Proceedings of 2021 IEEE
International Conference on Mechatronics and Automation
August 8 - 11, Takamatsu, Japan

A Novel Surgeon Training System for the Vascular Interventional Surgery based on Augmented Reality

Jian Guo[1], Shichen Jia[1]
Shuxiang Guo[1,2*]

[1]*Tianjin Key Laboratory for Control Theory & Applications in Complicated Systems and Intelligent Robot Laboratory Tianjin University of Technology*

Binshui Xidao 391, Tianjin, China

jianguo@tjut.edu.cn;
191417317@qq.com;

[2]*Department of Intelligent Mechanical Systems Engineering Faculty of Engineering Kagawa University*

2217-20, Hayashi-cho, Takamatsu 761-0396, Japan

*Corresponding author: guo@eng.kagawa-u.ac.jp

***Abstract* - At present, cerebrovascular disease has become the number one killer of human health, and minimally invasive interventional surgery is the main treatment. In order to solve the problem of radiation to doctors, master-slave surgical robot is favored. Even so, vascular interventional surgery has high requirements on the operation of doctors, and the actual modeling and training costs are high. The research of vascular interventional surgeon training system based on virtual reality technology is becoming more and more important. The quality of a training system depends on whether it can provide a realistic training environment. In the past, most of the researches mainly focused on the reality of touch, such as the simulation of the force on the guide wire in the blood vessel, but less on the human-computer interaction. A good human-computer interaction can also improve the doctor's operation experience and effect. To solve this problem, this paper built a training system based on augmented reality technology, and proposed gesture interaction to replace the traditional interaction. Through experiments, the system can satisfy the greater operational needs of doctors and improve the training effect.**

***Index Terms* - *Vascular interventional surgery, training system, augmented reality, gesture interaction*.**

I. INTRODUCTION

With the continuous development of social economy and the continuous improvement of people's aquatic products, vascular diseases incidence rate is higher in middle-aged and old people in China[1]. At present, the most direct and effective method for vascular diseases is minimally invasive vascula interventional surgery[2]. Minimally invasive vascular Interventional surgery refers to the process of doctors inserting the catheter into the patient's blood vessels through the movement and rotation of the catheter before and after the catheter, and moving in the vessel until the patient focuses, thus achieving the treatment of related diseases. Compared with traditional operation, minimally invasive vascular interventional surgery can significantly reduce the pain of patients because of its advantages of small injury, less pain and faster recovery. Therefore, it has been widely used.

Because of the still existing difficulties, the current level of interventional surgery technology and the patients' needs do not match, resulting in the low rescue rate of vascular diseases, high mortality. There are three main problems in the existing vascular interventional surgery: first, doctors operate under long-term radiation conditions, which seriously endangers the health of doctors [3]; secondly, doctors need to perform a minimally invasive intervention operation for several hours, and they also need to wear heavy lead clothing, which will cause doctors' extreme fatigue and affect the accuracy of the operation.With the support of modern science and technology, the training system of vascular interventional surgeons based on virtual reality can solve this problem well. It can let doctors carry out surgery training in the environment without radiation, and ensure the doctor's health to the maximum extent. At the same time, the application of virtual reality doctor training system can save costs, improve training effect and avoid some ethical problems[4].

But virtual reality still has its limitations. When doctors train in virtual system, because of the large difference between the virtual environment and the real world, they may reduce the sense of reality visually. To simulate a real operation, a lot of fine modeling is needed. Even if VR(Virtual Reality)glasses greatly improve the immersion, it will bring about some negative effects such as vertigo, and VR is totally closed, which limits its interactive function. Generally, it needs to connect external instruments. AR(augmented reality) technology developed at the same time with VR is favored, and provides another kind of visual training.

At the same time of improving visual interaction, human-computer interaction is also a very important link. There is a gap between the training system using mouse or other peripherals and doctors. The research shows that the use of gesture interaction can greatly improve the user's learning experience, and intuitively improve the continuity of training. Therefore, the introduction of gesture interaction is also necessary.This paper built a training system based on augmented reality technology, and proposed gesture interaction to replace the traditional interaction.

This paper is divided into five parts, the introduction of the current status of interventional surgery, the introduction of the work of the augmented reality system, the construction of the interact system , the display of the final effect, and the conclusion.

II. OVERVIEW OF PLATFORM

The master-slave manipulator is the Master manipulator operated by the doctor,which will transmits the doctor's surgical actions to the Slave side or training system through communication[5].After receiving the doctor's action information sent by the Master side, the client implements the corresponding action.

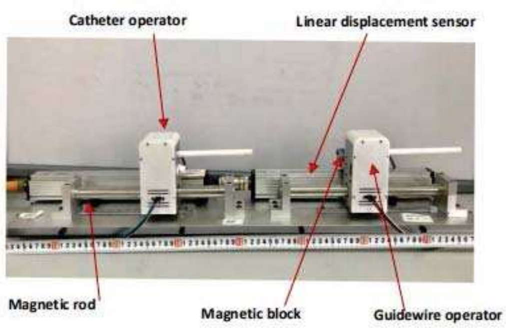


Fig .1 Master manipulator

The training system cooperated with the Mater manipulator aims to improve the training proficiency of doctors, as shown in Fig. 2 After receiving the signal from the Mater manipulator, the guide wire in the training system will move forward with the action of the Mater manipulator, and the doctor can feel the force feedback in real time and observe the movement state of the guide wire.

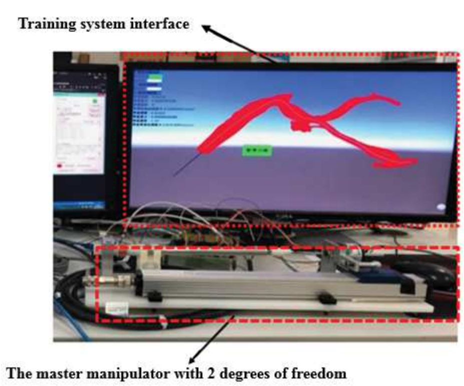


Fig .2 Training system

Augmented reality technology(AR), which is different from VR technology. AR technology emphasizes the interaction with the real world. By collecting images from the real world, combining them with virtual objects after computer processing, and displaying them through terminals (computer screen, mobile phone, headgear, etc.), people can get a real and more informative experience.

Augmented reality is the product of computer vision, the real scene is recorded to the screen (or program) through the lens, the scene in the screen is changed through image processing, and the image information is obtained through the screen to change the world in our eyes[6].

A. *Coordinate Technology between Virtual Model and Real Environment*

The main technology of AR system is three-dimensional registration, that is, spatial registration based on the real world. After collecting the real world information, how to make the virtual object appear in the designated position is a problem. This paper uses AR technology based on image recognition, takes the recognized image as the reference system, combines the real coordinate system with the virtual coordinate system, and finally determines the location of the virtual object location.

If the sitting posture of the virtual object in the world coordinate system is t, the posture in the camera coordinate system is TC, and the external parameter matrix of the camera is TG, the physical meaning of the matrix represents the posture of the camera in the world coordinate system, then the transformation relationship between the virtual object in the world coordinate system and the posture in the camera coordinate system is:

$$T_C = T_M^C T_M \tag{1}$$

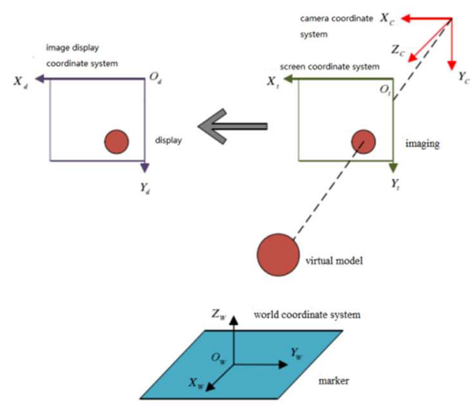


Fig. 3 Principle of the 3D registration

The imaging plane coordinate system is used to image the scene taken by the camera, and is parallel to the x.oy of the camera coordinate system. The imaging position (x, y,) of the object can be deduced from the similar triangle. Assuming that the position coordinates of the object in the camera coordinate system are (XC, YC, ZC), the specific value can be determined by the pose matrix T in the camera coordinate system. It is concluded that if the imaging focal length is f, then the derivation formula of the imaging position of the object is as follows:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} f & 0 \\ 0 & f \end{bmatrix} \begin{bmatrix} x_c \\ y_c \end{bmatrix} \tag{2}$$

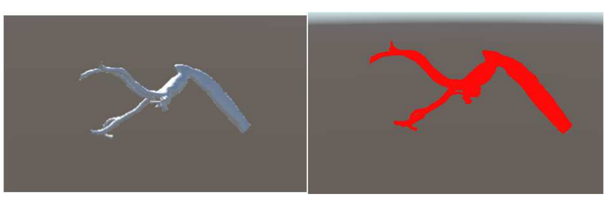
The position of an object in the imaging plane displayed on the display screen depends on the resolution of the screen on the one hand and the display range on the other. Suppose that the content displayed by the coordinate origin of the display corresponds to the point (m, n) of the imaging plane coordinate system, R and S represent the pixel value of unit length in the horizontal and vertical directions, that is, the measurement index of resolution. The relationship between the position (x, y) of the object from the imaging plane coordinate system and the position (x, y) of the image display coordinate system is as follows:

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & s \end{bmatrix} \begin{bmatrix} x_t - m \\ y_t - n \end{bmatrix} \tag{3}$$

The more complex the image is, the better the recognition effect will be. This paper selects the back of the author's ID card as shown in the Fig. below. Chapter4 shows the recognition results given by vuforia official website, which shows the main areas of image recognition and lays the foundation for the establishment of virtual keys below.

B. *Model Display based on Pattern Recognition*

After the location problem is solved, it is the selection of the object. In this paper, a segment of blood vessel is selected as the display object, and the model is completed by the laboratory. Using Mimics software to cut the CTA(CT angiography) data set, extract the blood vessel features, and complete the three-dimensional reconstruction[5].Then a segment of blood vessel with better effect is selected as the training object. As shown in the Fig. 4.



(a).blood model imported (b).blood model after rendering

Fig. 4 Vessel model in unity

This is the first cut blood vessel model, which is imported into unity for further rendering. Then the catheter and guide wire are simulated by modeling software, which are also imported into unity. Based on the completed work in the laboratory, the interaction program of the three is written. Finally, the AR system is completed by using the technology provided by vuforia. The final results are shown in Chapter 4.

C. *Model Interaction*

After modeling the blood vessel, we need to model the guide wire and the catheter and make them interact with each other in order to simulate a complete operation. The laboratory used Maya software to complete the modeling work, detected the collision through the bounding box method, and realized the simulation of various resistances in the blood vessel in the programming. In order to help doctors improve the training effect, the laboratory also put forward some suggestions A path planning algorithm based on scene projection matrix, the final effect is shown in the Fig.5, our ultimate goal is to combine the model with AR system, so that other modeling work can be saved, but still better visual effect can be achieved.

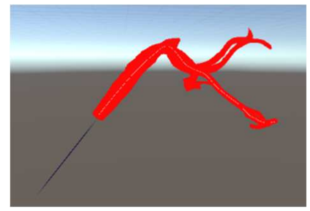


Fig. 5 Movement of guidewire in vessel

C. *AR System Construction*

The existing training system is built in unity. On this basis, we introduce the vofuria engine to realize the rendering of virtual scenes. In order to obtain the real world information, we must have the corresponding tools. As mentioned above, we need the camera to act as the "eye" of the system to get the real scene. Selecting HD camera can ensure the captured picture Clarity, the virtual model we have built is to integrate the virtual model with the real scene through the processing of vofuria engine. In order to accurately calibrate the location of the virtual model, we adopt the three-dimensional registration method based on the identifier to build the world coordinate system based on the identifier to integrate other coordinate systems together.

III. Interactive Function Based On Gestur Recognition

In the process of training, in addition to surgery, doctors may have other human-computer interaction needs, such as pause training to seek rest, enlarge the model to facilitate observation, and rotate the model to find different information. The realization of these functions can also enhance the training effect of doctors. In the traditional interaction, most of the systems are through the mouse touch screen for doctors, it will destroy the continuity of training, and even affect the touch. This paper proposes a gesture recognition technology based on normal camera, which uses the doctor's own gesture to replace the control of other external devices to meet the diversified experience.And considering that there is a sampling interval when the doctor operates the main terminal, the error will accumulate after a long time, so we can use gesture operation to manually adjust. Ensure the accuracy of training.

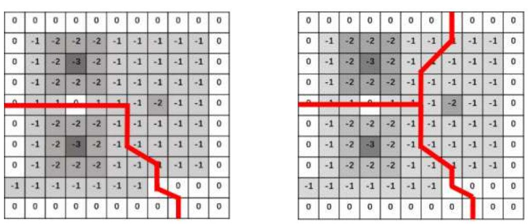
A. Skin color Detection

First of all, we need to make it clear that skin color is a human feature, and its distribution is relatively consistent in the overall and appearance of the hand. Using this phenomenon, we can effectively detect the naked human organs[7]. Skin color detection is to segment the skin color region in the image, which is essentially to classify all the pixels in the image. They belong to skin color category or non skin color category.In this paper, skin color model based on H-CB-CR is used to detect skin color to complete subsequent cutting.

H-CB-CR model is a hybrid model of HSV model and YCbCr model. The H component of the former and the CB and Cr components of the latter are selected respectively. Skin color segmentation based on H value can better retain skin color pixels, but it is not detected.There are also more non skin color pixels filtered. Skin color segmentation based on CBCR value can filter out non skin color pixels better, but in high light, skin color pixels lose too much[8].

B. *Image Segmentation*

Image segmentation is the second step of gesture recognition. Its purpose is to segment the connected area of skin color and obtain its location and boundary information. Location information directly determines the success of gesture tracking. Accurate extraction of skin color region can improve the processing speed and anti-interference ability of the next gesture recognition algorithm[8]. This paper uses a watershed algorithm based on marker t,seed filling method to obtain the boundary information and location information of skin color connected region. In order to ensure that gesture recognition is only carried out in the skin color connected area.



(a) Watershed algorithm with marker (b)Watershed algorithm

Fig. 6 Comparison between two algorithms

Seed filling algorithm is a region based image segmentation method. It needs to give image data and a pixel point in the target area. The single pixel point is regarded as the seed, and then it spreads and merges from the seed gradually to form the required segmentation area. In gesture recognition, the pixel value of the contour of the image cut by watershed is 0, so the point with pixel value of 0 can be selected as the seed and the eight connected search algorithm can be used.

C. *Gesture Tracking and Recognition*

The area of gesture recognition task has been determined, and the next step is to recognize the gestures captured in this area. This paper provides five recognizable gestures, which are numbers one to five represented by hand. The height width SR model is used to track the gesture area, and the gesture recognition is carried out through template matching.When the recognition gesture and the template reach a certain coincidence, the system outputs the corresponding recognition results, otherwise there is no output.

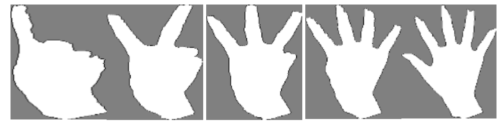


Fig. 7 Template of the gesture

D. *Construction of Interactive System*

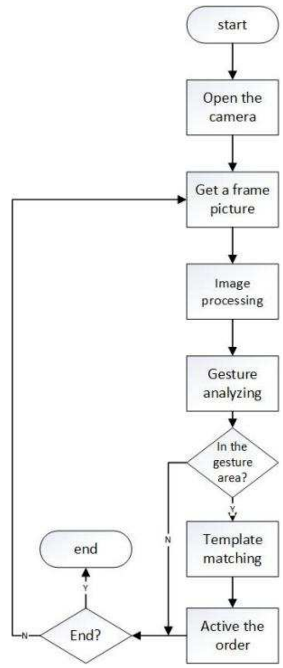


Fig. 8 Gesture interaction flow

When the hand gesture recognition finished, we need to add the interactive function. In this paper, the five recognized hand gestures are matched with mouse commands to complete the interactive function of the training system. The training system of our laboratory is connected with the main terminal independently developed by our laboratory to control the movement of the catheter/guide wire. The interactive system proposed in this paper greatly enriches the interactive function of the training system and meets the requirements of the medical laboratory we need more.

This system can recognize five kinds of gestures, namely numbers one to five, and each gesture can represent a command. In this paper, we choose to use gestures to simulate mouse operation. Different gesture combinations can achieve different operations, which is very suitable for the current popular graphical operation.

IV. EXPERIMENTS AND ASSESSMENTS

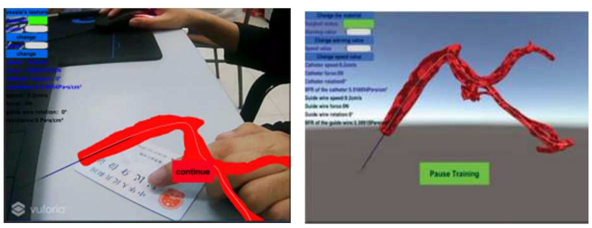
Experimental environment: unity + vuforia

Both augmented reality system and gesture interaction system need cameras to collect data, and both sides have certain requirements for angle, so two cameras are selected to serve different systems.

Unity is a real-time 3D interactive content creation and operation platform. Vufora is a mature augmented reality development platform, which has been built into unity. This paper uses this platform to build the basic framework of the training system. The experiment is divided into three parts, namely, the implementation of augmented reality, the implementation of hand interaction and the comparison of the training effect between the new system and the old system.

A. *AR System*

As shown in the Fig.9, When the marker appears in the field of vision of the camera, the virtual soft tissue also appears. In the experiment, we can clearly see the movement of catheter (blue) and guide wire (green) in the blood vessel, and the path (white) based on the laboratory work planning. Doctors operate the guide wire catheter with the path as the reference in the training.



(a) .AR system (b).VR system

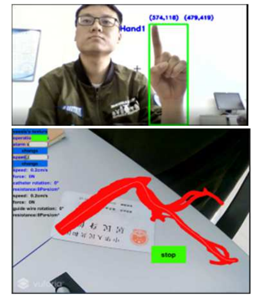
Fig. 9 Comparison between two systems

Compared with VR system,AR system not only inherits the characteristics of high fidelity in the traditional training process, but also has the advantages of low cost and repeatable training in virtual reality training system.

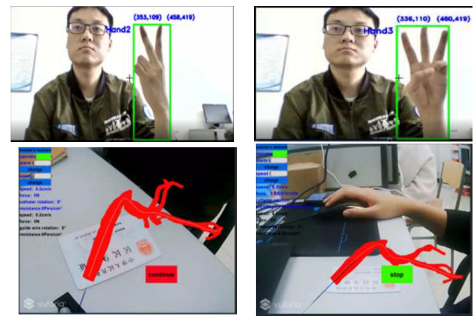
B. *Gesture Interaction System*

It is worth noting that gesture itself can represent a lot of instructions. In order to improve the operation range of gesture, and considering the characteristics of the training system in the laboratory, this paper combines the identified five gestures with windows commands to simulate the operation of the mouse. For different systems, gestures can express different commands.

As shown in Fig.10, different gestures are used to stimulate different system functions. Gesture "1" controls the start and pause of the system; The gesture "2" controls the rotation of the whole model. The rotation of the model follows the movement direction of the hand. It rotates counterclockwise to the left, clockwise to the right, and up and down to rotate up and down; The gesture "3" controls the movement of the catheter. The left motion drives the catheter to move backward, the right motion drives the catheter to move forward.When the gesture moves directly above or below, the catheter will rotate with the motion of the gesture, and the rotation angle changes.



(a)Gesture '1' makes system stop



(b)Gesture '2' makes model rotation (c) Gestue'3' controls the catheter movement

Fig. 10 Gesture interaction show

C. *Comparison of Training Effect between AR and VR System*

In order to ensure the authenticity and effectiveness of the experiment, we selected 6 healthy adult operators into three groups, two in each group, simulating the situation of doctors facing the training system. Each group of members operated five times on two training systems. After that, the training system was exchanged, the starting point and end point were reset, and then operated for 5 times again. Making the model rotate when reaching the given point. At the same time, observing the training status of each operator and records the training time.The Result shows that new system works well. Finally,we summarize the recorded data of the two systems and make the average value, and draw the curve as shown in the Fig. 12 below.

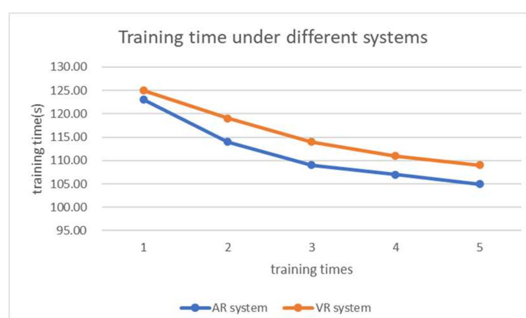


Fig.11 Training time under different systems

As the Fig. 12 shown, the blue curve is the training time in AR system with gesture recognition and the orange curve is the training time in VR system.The ordinate is the average of training time, and the abscissa is the training times. Through the observation of the operator, it can be seen that the operator can enter the state faster in the augmented reality system, that is, the training time in the initial stage of training is significantly faster than that in the virtual reality system, and the interaction time in the gesture operation is significantly shorter, which makes the training more fluent.

V.CONCLUSIONS

In view of some shortcomings of VR system, this paper proposed a method to apply augmented reality technology to the training system, which could reduce the workload of modeling and give doctors a more realistic and familiar visual experience. Combined with gesture recognition technology, a new interactive way was proposed to meet the diversified operation needs of doctors.The technology also allows doctors to manually calibrate possible errors and ordinary cameras could be used, which reduce the threshold of gesture interaction.This is just a preliminary framework. We still have a lot to improve, such as transplanting it to mobile devices such as smart glasses,using deep learning technology to make gesture recognition more accurate, building more complex models, and enabling doctors to complete more complex and comprehensive surgical training.

ACKNOWLEDGMENT

This research is supported by National Natural Science Foundation of China (61703305), Key Research Program of the Natural Science Foundation of Tianjin (18JCZDJC38500)and Innovative Cooperation Project of Tianjin Scientific and Technological (18PTZWHZ00090).

REFERENCES

[1] S. Guo, Q. Zhan, J. Guo, C. Meng, and X. Jin, "Vascular Environment Modeling and Verification for The Virtual Vessel Interventional Surgery Training System*", International Conference on Mechatronics and Automation*, pp. 1087-1092, 2018.

[2] J. Guo, L. Qi, S. Guo, C. Meng, and Q. Zhan, "Study on Tracking Stability for a Master-Slave Vascular Interventional Robotic System", *International Conference on Mechatronics and Automation*, pp.1387-1392, 2019.

[3] L. Zhang, S. Guo, H. Yu, and Y. Song, "Performance evaluation of a straingauge force sensor for a haptic robot-assisted catheter operating system," *Microsystem Technologies*, vol. 23, no. 10, pp. 5041-5050,2017

[4] Y. Wang, S. Guo, Y. Li, T. Tamiya, and Y. Song, "Design and evaluation of safety operation VR training system for robotic catheter surgery", *Medical & Biological Engineering & Computing*, vol.56, no.1, pp.25- 35, 2017.

[5] J. Guo, Y. Cheng, and S. Guo, "Three-dimensional reconstruction of brain blood vessels and algorithm implementation", *Journal of Harbin Engineering University*, vol. 40, no. 4, pp. 1-6, 2019.

[6] P. Milgram and A. F. Kishino, "Taxonomy of mixed reality visual displays." *IEICE Transactions on Information and Systems*, pp. 1321–1329, 1994.

[7] M. Cote, P. Payeur, and G. Comeau, "Comparative Study of Adaptive Segmentation Techniques for Gesture Analysis in Unconstrained Environments." *IEEE International Workshop on Imagining Systems & Techniques* IEEE, 2006.

[8] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," Proceedings. I*EEE Computer Society Conference on Computer Vision and Pattern Recognitio*, pp. 274-280 Vol. 1,1999.

[9] X. Yang, H. Wang, and Z. Xu, "Calibration and operation of a positioning robot used for minimally invasive vascular interventional surgery", *Progress in Modern Biomedicine*, pp.1-13, 2013.

[10]L. Zhang, S. Parrini, C. Freschi, V. Ferrari, S. Condino, M. Ferrari and D. Caramella, "3D ultrasound centerline tracking of abdominal vessels for endovascular navigation", *International journal of computer assisted radiology and surgery*, vol. 9, no.1, pp.127-135, 2014.

[11]J. Guo, X. Jin, and S. Guo, "Study of the Operational Safety of a Vascular Interventional Surgical Robotic System", *Micromachines*, vol. 9, no. 3, 2018.

[12] Y. Wang, S. Guo, Y. Li, T. Tamiya and Y. Song, "Design and evaluation of safety operation VR training system for robotic catheter surgery*",Medical & Biological Engineering & Computing*, vol.56, no.1, pp.2535, 2017.

[13]N. Rudarakanchana, I. Van Herzeele, L. Desender, and N. JW Cheshire, "Virtual reality simulation for the optimization of endovascular procedures: current perspectives", *Vascular Health and Risk Management*, pp.195-202, 2015.

[14]S. Wijewickrema, B. Copson, Y. Zhou, X. Ma, R. Briggs, J. Bailey, G. Kennedy and S. O'Leary, "Design and Evaluation of a Virtual Reality Simulation Module for Training Advanced Temporal Bone Surgery", *IEEE International Symposium on Computer-based Medical Systems*, pp.7-12, 2017.

[15] S.Nicolau, L.Soler, and J.Marescaux, "Augmented reality in laparoscopic surgical oncology", *Surgical Oncology*. vol. 20,no. 3,pp.189-201,2011.

[16]A.K. Bourke, J.V. O'Brien, and G.M. Lyons, "Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm",*Gai&Posture*,vol.26,no.2,pp.194–199.,2007.